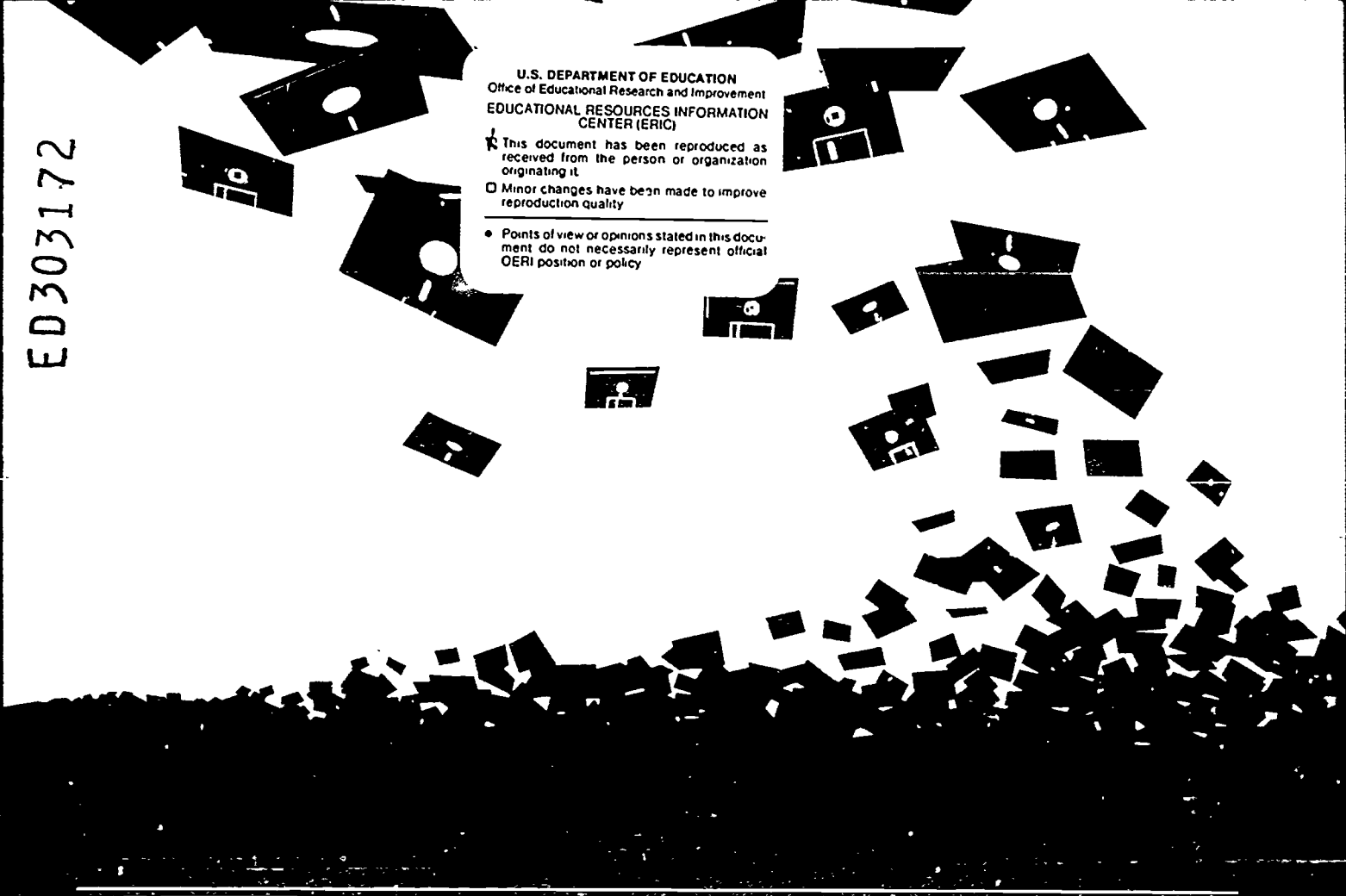
DOCUMENT RESUME

ED 303 172 IR 052 622

| | |
|---|---|
| TITLE | A Strategic Plan for Managing and Preserving Electronic Records in New York State Government: Final Report of the Special Media Records Project. |
| INSTITUTION | New York State Education Dept., Albany. State Archives and Records Administration. |
| PUB DATE | Aug 88 |
| NOTE | 44p. |
| PUB TYPE | Reports - Descriptive (141) |
| EDRS PRICE | MF01/PC02 Plus Postage. |
| DESCRIPTORS | *Archives; Information Technology; *Recordkeeping; *State Agencies; *State Programs |
| IDENTIFIERS | *Machine Readable Data; *New York; Records Management |

ABSTRACT

This report presents the final recommendations of the Special Media Records Project for the improved management and preservation of machine-readable records in New York State government agencies. The first of three parts of the report discusses strategic issues in the management and preservation of electronic records, including key trends in information technology; the relationship between paper and electronic records; and the organizational context. Objectives for improved management and preservation of electronic records are outlined in the second part with attention to records analysis and disposition, agency services, and archival services. The third part describes specific activities that must be undertaken during the next five years to integrate machine-readable records into state government records management programs, and to develop the capacity at the State Archives to select, acquire, and preserve records with enduring value in machine-readable form. An executive summary and a table depicting resource needs for these activities over a 5-year period are included. (MES)



# A STRATEGIC PLAN FOR MANAGING AND PRESERVING ELECTRONIC RECORDS IN



NEW YORK
THE STATE OF LEARNING

THE UNIVERSITY OF THE STATE OF NEW YORK
THE STATE EDUCATION DEPARTMENT
THE STATE ARCHIVES AND RECORDS ADMINISTRATION

# A STRATEGIC PLAN FOR MANAGING AND PRESERVING ELECTRONIC RECORDS IN NEW YORK STATE GOVERNMENT:

## Final Report of the Special Media Records Project

The University of the State of New York
The State Education Department
State Archives and Records Administration
Albany, New York

AUGUST 1988

## THE UNIVERSITY OF THE STATE OF NEW YORK

### Regents of The University

| | |
|---|---|
| MARTIN C. BARELL, Chancellor, B.A., I.A., LL.B. | Muttontown |
| R. CARLOS CARBALLADA, Vice Chancellor, B.S. | Rochester |
| WILLARD A. GENRICH, LL.B. | Buffalo |
| EMLYN I. GRIFFITH, A.B., J.D | Rome |
| JORGE L. BATISTA, B.A., J.D. | Bronx |
| LAURA BRADLEY CHODOS, B.A., M.A. | Vischer Ferry |
| LOUISE P. MATTEONI, B.A., M.A., Ph.D. | Bayside |
| J. EDWARD MEYER, B.A., LL.B. | Chappaqua |
| FLOYD S. LINTON, A.B., M.A., M.P.A. | Miller Place |
| MIMI LEVIN LIEBER, B.A., M.A. | Manhattan |
| SHIRLEY C. BROWN, A., M.A., Ph.D. | Albany |
| NORMA GLUCK, B.A., M.S.W. | Manhattan |
| JAMES W. MCCABE, Sr., A.B., M.A. | Johnson City |
| ADELAIDE L. SANFORD, B.A., M.A., P.D. | Hollis |
| WALTER COOPER, B.A., Ph.D. | Rochester |
| GERALD J. LUSTIG, B.A., M.D. | Staten Island |

**President of The University and Commissioner of Education**
THOMAS SOBOL

**Executive Deputy Commissioner of Education**
THOMAS E. SHELDON

**Deputy Commissioner for Cultural Education**
CAROLE F. HUXLEY

**State Archivist and Records Administrator**
LARRY J. HACKMAN

# TABLE OF CONTENTS

# EXECUTIVE SUMMARY

This report presents the final recommendations of the Special Media Records Project for improved management and preservation of machine-readable records in New York State government agencies. In 1985, the State Archives, in cooperation with the Governor's Office of Management and Productivity and 19 State agencies, initiated the Special Media Records Project. The purpose of this two-year project was to assess the adequacy of State government policies and procedures for the management of computer-generated, machine-readable records, and to develop a program for the long-term preservation of selected, valuable machine-readable records at the State Archives.

Computers are used to create, process, and store an increasing percentage of the State's public records. Much of the information that State agencies used to store solely on paper, now exists in electronic form on magnetic disks, tapes, and floppy disks. The use of automated systems can bring tremendous benefits to State government because information in automated systems is easier to access, manipulate, and disseminate. Yet, the use of computers for record keeping also creates problems for State government. Computers have contributed to an information explosion which has inundated some program managers with more information than they can handle.

Even though an important part of the public record exists in machine-readable form, computer-generated records were not integrated into New York State's centralized records management programs. Comprehensive records management programs include investigation into the conditions, quantity, and nature of records in State agencies, study of records keeping methods, provision for transfer of inactive records to the State Records Center, authorization for the disposal or destruction of records, and preservation of records with enduring value at the State Archives. Records management programs are only partially developed for hard copy (paper and microfilm) records, and few programs address the special problems of managing machine-readable records.

The lack of attention to the management, retention, and selective preservation of machine-readable records creates numerous problems for data processing departments, program staff, and researchers. Even though technology is available to enhance data sharing and make information more accessible, the lack of information about the content, condition, and location of electronic data processing (EDP) records makes access difficult. Without systematic review of data retention requirements, some machine-readable records are maintained longer than necessary to meet the information needs of agencies. Others, which may have secondary uses or enduring value, are erased or destroyed without the review or authorization required by law.[1] The need for better management of machine-readable records is especially acute because these records are stored on media which will deteriorate without active

[1]Section 57.05, Subdivision 11 of the Arts and Cultural Affairs Law (formerly Sec. 186 of the State Finance Law) requires review by the Attorney General and the Comptroller, and approval by the Commissioner of Education prior to the destruction of any public records. Public records are generally defined as any material created or received by a government agency regardless of physical form or media.

intervention to increase the longevity of the records.

The use of computers for record keeping is also having a profound impact on the ability of the State Archives to preserve State records with enduring value. According to a recent report by the Committee on the Records of Government, the United States is in danger of losing its documentary heritage because federal, state and local governments are shifting rapidly to electronic record keeping without making provisions to identify, select, and preserve valuable records in machine-readable form.[2] The situation in New York State is similar to the national picture. Archivists have developed procedures and techniques to transfer valuable machine-readable records to repositories and to prolong the life span of these records.[3] However, better management of machine-readable records while they remain in agency custody and increased resources for archival programs are essential in order to preserve selected, valuable machine-readable records.

It was assumed from the outset — and this report confirms — that the archival preservation of machine-readable records is dependent upon the quality of records and information management programs in State agencies. However, long-term preservation of selected machine-readable records is not the only reason to improve the management of EDP information resources. Policies and procedures for the orderly retention and disposition of machine-readable records are essential to ensure that adequate documentation of government activities is maintained, to protect the legal and fiscal interests of the State and its citizens, and to provide for the efficient administration of government programs. Therefore, the recommendations in this report address not only the essential needs for archival program development, but also include areas where government-wide initiatives are needed and where State agencies must assume more responsibility for managing their machine-readable information resources.

When the Special Media Project began, responsibility for records management and archival programs was divided between the Bureau of Records Management in the Office of General Services and the State Archives in the State Education Department. Fundamental changes in the organization and administration of records management activities in New York State now provide the basis for a more comprehensive approach to the management and selective preservation of machine-readable records. As of April 21, 1987, statutory responsibility for Executive Branch records management programs was transferred to the State Education Department and combined with the archives program. An expanded records management function is being developed in the Education Department by the State Archives and Records Administration (SARA). One goal of this program is to integrate machine-readable records issues into records analysis, retention scheduling, and archival preservation.

This report addresses the archival and records management problems of electronic records in the form of a five-year strategic plan. The report is divided into three sections. Statewide records management and archival programs which accommodate modern information technologies must respond to dynamic changes in the ways records are created, maintained, and used. Therefore, the first section of the report discusses the impact of key trends in information technology on the capacity of State agencies and SARA to manage and selectively preserve records to meet legal, fiscal, administrative, and research needs. The record keeping implications of five important developments are analyzed: 1) the increased use

[2]Committee on the Records of Government, *Report* (Washington, D.C., March 1985).

[3]Margaret Hedstrom, *Archives & Manuscripts. Machine-Readable Records* (Chicago. Society of American Archivists, 1985).

of computers for record keeping, 2) wider distribution of information and records, 3) decentralization, 4) increasingly diverse applications, and 5) the use of information systems design methodologies. This section also examines the relationship between computerized and hard copy records and the changing organizational context in which records are created and managed.

The second section of the report discusses seven broad objectives for records management and archival programs which will help SARA and State agencies respond to the challenges posed by new record keeping technology. During the next five years, SARA will work with agencies and with government-wide advisory bodies to develop guidelines for the identification, description, retention and selective preservation of machine-readable records. Implementation of such programs should ensure that government programs are adequately documented, and that information in machine-readable form is accessible (with appropriate safeguards for confidential data) to State government and the public. SARA will also develop programs in education, training, technical assistance, and will improve off-site storage facilities to help agencies better manage EDP records. To meet its responsibility to preserve the historically-valuable records of State government, the State Archives will develop its capacity to select, accession, describe, preserve and make available to researchers machine-readable records with enduring value.

The final section of the report outlines in detail 27 activities that State agencies and SARA must undertake to accomplish better management of machine-readable records during the next five years. Each activity and each objective is linked to a SARA Bureau in its State Government Records Division: Records Analysis and Disposition, Agency Services, and Archival Services. For each activity a time frame and brief explanation are included.

To accomplish the objectives set out in this plan, SARA will utilize existing resources and will require additional resources. SARA will request funding for two additional archivists, beginning in FY 1989-90, one in Records Analysis and Disposition and one in the Collections Management Unit, Bureau of Archival Services. These positions are needed to establish guidelines for retention of EDP records in machine-readable form, to evaluate EDP records that have potential archival value, and to acquire and preserve selected valuable EDP records. SARA will request funding for one additional Public Record Analyst in FY 1990-91 to develop programs for managing records from automated office information systems. The equivalent of approximately 1 to 1.5 Full Time Equivalents (FTE) of existing staff will also be devoted annually to management and archival preservation of EDP records. Throughout the period covered by this plan, approximately $27,000 is needed each year for contract data processing services and supplies, and for bringing experts in information technology to Albany to present workshops and seminars. In FY 1989-90, SARA will request $150,000 for installation of adequate tape storage vaults, and $60,000 to begin a study of optical disk technology for archival preservation. Additional resources may be needed, beginning in FY 1990-91 to appraise and accession valuable data that is in immediate danger of loss or destruction, to make possible capital improvements to the Records Center tape storage vaults, and to start a tape maintenance program. Resource requirements for these projects cannot be estimated until additional assessments are completed during FY 1988-89 and 1989-90.

# INTRODUCTION

**Background.** Several recent studies have stressed the inadequacy of records management in New York State and proposed new initiatives to deal with this problem.[4] Few agencies have developed programs to manage their paper or machine-readable records because of inadequate training, absence of centralized direction, and insufficient resources. In most agencies, there is no centralized source of information about which records exist, where they are located, or how long they need to be kept. Records and data in both paper and machine-readable formats may be retained longer than necessary for agency administrative purposes or for research use by parties outside the agency, while records of long-term value may be destroyed prematurely.

The lack of centralized information about records makes it difficult for users to locate records and data that they could use for research or administrative purposes. Agencies often collect and store redundant data and records resulting in a duplication of effort and use of expensive storage space. Without better records management it is also difficult for archivists to identify those records which should be appraised for their long-term value. Less than one-third of all government records have been effectively administered under the existing records management provisions, and with few exceptions, machine-readable records have been completely neglected. As a consequence, the retention, disposition, and selective preservation of most government records, and especially of EDP records, are ad hoc and haphazard activities.

In 1984, the State Historical Records Advisory Board recommended to the Governor that a comprehensive records management program be developed for State government records. In recognition of the widespread application of new information technology for records keeping, the Board recommended that special attention be devoted to the management and preservation needs of audio-visual and electronic records. In 1985, the State Archives, in cooperation with the Governor's Office of Management and Productivity, began the Special Media Records Project — a two-year study of computer and audio-visual records in State agencies.

The first stage of the Special Media Records Project was an extensive survey of the nature, quantity, and condition of special media records in 19 selected New York State agencies. The initial findings and recommendations of the project were reported in a preliminary report, "Computer and Audio-Visual Records in New York State Government," April, 1986.[5] The preliminary report recommended that the State Archives should begin a systematic program to identify and appraise machine-readable records for enduring value in State agencies, that the Archives and agencies work together to establish programs and policies for their management, retention, and

[4]New York State Historical Records Avisory Board, *Toward a Usable Past*, Albany, 1984, Governor's Office of Management and Productivity and New York State Education Department, State Archives, *Computer and Audio-Visual Records in State Government*, Preliminary Report of the Special Media Records Project, Albany, 1986, and Governor's Office of Management and Productivity, "Consolidating Archival and Records Management Function in New York State: A Plan for the Future," Albany, 1986.

[5]Copies of this report are available from the New York State Archives.

archival preservation, and that the Archives obtain the resources and facilities needed to enhance its own ability to preserve these records.

In 1985 — about the same time that the Special Media Records Project began — Governor Cuomo issued an Executive Order which required agencies to develop records management improvement plans and retention schedules for all records, including machine-readable records. The Governor's Office of Management and Productivity is coordinating and providing oversight for this initiative, the Governor's Records Management Improvement Project, by sponsoring records management training, requiring agencies to inventory all records, and reviewing records management plans from Executive Branch agencies. Following an extensive study of the State's records management and archival programs by the Office of Management and Productivity, the Legislature recently consolidated these two programs and transferred responsibility for centralized records management services from the Office of General Services to the State Archives and Records Administration (SARA), in the State Education Department. This administrative change was intended to increase the centralized support and coordination needed for an ongoing records management program in New York State government.

**Scope of the Report.** This report presents the final recommendations of the Special Media Records Project for the improved management and selective preservation of computer-generated EDP records in New York State agencies.[6] The report recommends a series of activities to be carried out by SARA and by Executive Branch agencies during the next five years in order to improve the management of records created with new information technologies and to enhance the ability of the State Archives to preserve selected, valuable records in machine-readable form. Many of the recommendations in this report may also apply to the Legislative and Judicial branches of government and to local government records, but the problems of electronic record keeping in these governmental organizations need further study before precise recommendations can be developed.[7]

This report stops short of outlining a broad information policy for New York State government. A general information policy would address such issues as procurement of computer, software and telecommunication systems, long-range information systems planning, data sharing, access, privacy protection and security of computerized records, and the development of systems design methodologies. Although the project staff believes that such policies are essential to improved information management in New York State government, no single body can assume responsibility for all aspects of information management. It is incumbent upon government-wide policy and advisory bodies to develop a general information policy and to coordinate the numerous areas where these issues overlap. The recommendations in this report are limited to measures needed to improve the management and selective preservation of EDP records because these issues reflect the primary mission of the State Archives and Records Administration.

This report is divided into three sections. The first section of the report outlines strategic issues in the management and archival preservation of records which result from the increased use of new information technologies. The second part of the report discusses seven objectives for the improved management of machine-readable records and for the long-term preservation of selected valuable records at the State Archives.

[6]SARA has prepared a separate report on audio-visual records.

[7]SARA's Division of External Programs, Local Records Bureau provdies records retention guidelines and other services to local governments. For an initial assessment of EDP records issues in local governments, see New York Local Government Records Advisory Council, *The Quiet Revolution: Managing New York's Local Government Records in the Information Age* (Albany, N. Y., December 1, 1987).

The final section of the report describes 27 activities to be carried out by the State Archives and Records Administration, with assistance from Executive Branch agencies and information management policy and advisory bodies, during the next five years.

PART ONE:

# STRATEGIC ISSUES IN THE MANAGEMENT AND PRESERVATION OF EDP RECORDS

New York's records management and archival programs at present are directed primarily toward hard copy (paper and microfilm) records. The electronic age raises new issues, challenges and opportunities for improved delivery of information and for better management of the State's information resources. The potential for more efficient processing of information, improved access, easier retrieval, and cost-effective storage, however, will not simply result as a natural outcome of the use of new information systems.[8] Obtaining the potential benefits of the information revolution will be highly dependent on the ability of the State to manage its data and records. Such an initiative should start with a systematic analysis of existing information resources and an assessment of the needs of agencies to ensure that the data and records they need to carry out their missions are available.

Most legislation regulating the management, disposal, and preservation of public records was enacted in the precomputer age. Records management programs generally have neglected the access, retention, and preservation needs of machine-readable records. Since computers are used to produce, store and retrieve an increasing proportion of the State's records, the lack of programs to protect, manage, and selectively preserve this information will become increasingly critical. Machine-readable records must be incorporated into the existing laws, regulations, policies, and procedures governing the management of public records to ensure that all of the State's information resources will remain accessible to public officials and private citizens. This section of the report discusses key issues in the management and selected preservation of EDP records and proposes a strategic approach to their improved management during the next five years.

## 1.1 KEY TRENDS IN INFORMATION TECHNOLOGY

The increased use of computers, the implementation of office automation systems for text processing and electronic mail, and the development of advanced communications networks are symbolic of the information revolution. Computer technology, used initially to process numeric data, can now be coupled with sophisticated software and advanced telecommunications systems to process and distribute almost any type of information. The use of computers is no longer limited to the processing of numbers and small amounts of text. Increasingly, all forms of information including text, voice communications, maps, graphs, and drawings can be handled by advanced information technology. The expanded use of computers, telecommunications, and other information technologies is creating a growing pool of information with which program managers, records managers, and archivists must cope.

**Increasing use of computers in State government.** One of the most obvious key trends in information technology is the increasing use of computers for all forms of record keeping and

[8] In this report an information system is defined as a collection of computer hardware, software, procedures, and data organized to accomplish a special function or set of functions.

communications. The number of applications supported by centralized, mainframe computer systems is still increasing in spite of a parallel movement toward decentralized, microcomputer-based systems. The Special Media Records Project survey identified more than 385 mainframe-based production systems in 19 Executive Branch agencies. Based on this survey, it is likely that more than 1,000 centralized, automated information systems are created and maintained by State agencies. Preliminary results from an inventory of records in 15 agencies indicate that data processing is used to process and produce more than 20 percent of the paper and microfilm records in New York State government agencies.[9]

**Wider distribution of information in electronic form.** Expanded telecommunications capabilities are increasing the distribution of information in electronic form. In July 1987, New York State installed a new telecommunications system which allows transfer of voice or data communications over one fiber optics network. This new technology may also help to resolve some vital management issues for public sector administrators. According to a recent Sperry Corporation survey of public officials, one of three constant barriers to effective management in the public sector is the geographic dispersion of operating units. Nearly one-third of the respondents reported that they frequently use records and information from other government agencies, while nearly half sometimes rely on external sources of information.[10] Advanced telecommunications networks offer the potential to share information between different geographic locations of one agency and among different agencies.

**Decentralization.** The more recent development of microcomputers, word processors, and local area networks is rapidly increasing the extent to which program managers, rather than EDP professionals, create and manage machine-readable records. Office automation systems supply information for policy analysis, research, and evaluation in New York State agencies but there has been no systematic study of office automation or electronic mail in State government. The Sperry Corporation survey indicates how widespread office automation systems have become in public agencies generally. More than two-thirds of the 5,200 public sector officials surveyed use word processing technology, and more than 40 percent rely on computers for report generation.

It is unlikely that decentralized systems will replace the large production systems supported by mainframe computers. In the future, agencies' automated information resources will consist of a mixture of centralized and decentralized databases, files, and documents. However, the availability of low cost and user-friendly microcomputer and word processing systems has greatly accelerated the creation of records in machine-readable form. Many of these records are created, accessed, and controlled by staff in program units rather than central EDP units. Therefore, programs to improve the management of machine-readable records must reach beyond central EDP units to users in program offices.

**More diverse applications and software.** The increased sophistication of software is another key trend in information systems technology which has encouraged the expanded use of computers for records creation, processing, and storage. Most mainframe-based production systems rely on database management and other complex software to process data, manage editing and updates, handle on-line queries, and produce a variety of reports and other hard copy

[9]Data collected by the Governor's Records Management Improvement Project, a project sponsored by the Office of Management and Productivity to inventory all State records and develop records management improvement plans.

[10]Sperry Corporation, *Public Sector 1985 Survey Results*, p.2.

output. The results of the Special Media Project survey suggest that nearly three-quarters of the State's centralized EDP systems are designed for use with in-house software, database management systems, or statistical packages. Most microcomputer-based systems rely heavily on software packages for data entry, formatting, and report generation.

The use of sophisticated software has distinct advantages for data processing and information retrieval. In large production systems, database management systems reduce data redundancy and allow users to combine data in a variety of ways. User-friendly software packages for microcomputers and word processing systems enable individuals with limited training and virtually no knowledge of computer concepts or programming techniques to use computers.

The increasing use of sophisticated software, however, poses problems for data exchange, information management, and archival preservation of machine-readable records. The hardware and software systems that cannot communicate with each other create obstacles to the flow of information between systems within one agency, among agencies, and between agencies and the Archives. The use of data file structures which require special software complicate the archival preservation of EDP records. To preserve software-dependent data and keep it accessible, archives have to maintain the data and the software or rewrite the data into a format that does not require special software. Creating data files in a software-independent format increases preservation costs if such capabilities are not part of the system design. Moreover, the use of data structures which do not require software has increasing limitations as computers are used for a much wider variety of applications, such as text processing, graphics, mapping, and computer-assisted design. Software dependency will become a major obstacle to the transfer, management, and preservation of information in electronic form as more sophisticated systems are used.

**Use of systems design methodologies.** Another important trend is the increasing use of more precise design methodologies in the development, implementation, and management of automated systems. Studies of data processing have demonstrated frequently that information systems do not function well unless they serve real needs of administrators and program managers for better processing and retrieval of information. Especially in the development of large production systems, a detailed analysis of information needs precedes the development of new automated information systems. Such design methodologies usually specify the hardware and software requirements for the systems, data flows, data base management, and the production of output reports.

Unfortunately, analysis and planning for the long-term management of the machine-readable and hard copy records produced by automated systems rarely is addressed in the system design methodology. Decisions about the disposition of machine-readable records often are not made until a system has reached its capacity in terms of disk or tape storage space. Although retention cycles usually are established for the input, output, edit, and other temporary data files, there has been little systematic analysis of the retention and disposition requirements for master data files, databases, and related hard copy documents. Better control, management, and preservation of EDP records require systematic review of the data to be created and stored by automated systems as part of the design methodology.

These trends taken together suggest that the future creation, storage, distribution and use of public records is likely to differ considerably from the traditional paper-based information systems to which most records managers and archivists are accustomed. Accompanying the spread of computer technology is a growing diversity of applications ranging from large-scale, centralized production systems, to automated text and

report generation, to computer-assisted design and decision support systems. Increasingly, decentralized and user-driven systems which support a wide variety of record keeping applications will exist along side of large, centralized databases.

### 1.2 RELATIONSHIPS BETWEEN PAPER AND ELECTRONIC RECORDS

**Overlapping content and functions.** Some experts predict that the "paperless office" will be commonplace by the 1990s as advanced information systems provide instantaneous access to information through computers, telecommunications, and optical disk systems. To date, however, few automated systems have eliminated the use of hard copy documents, and in many cases the use of computers has actually accelerated the creation of paper records. As a consequence, the contents and functions of machine-readable records and hard copy documents usually are closely related. Data may be extracted from a database to produce summary reports, printouts may be produced as a database is updated, or the correspondence stored on a diskette may also exist in an agency's files. For these reasons, data and records management must be coordinated to account for the closely related content and functions of machine-readable records and hard copy records. A comprehensive information management program must focus first on analysis of the information in records and then on control of the medium on which the information is stored.

**Special characteristics of EDP records.** EDP records differ in a number of ways from the paper and microfilm records. In the past, records management has focused on managing the physical entities on which information is recorded, such as files of paper documents. Information in machine-readable form, however can be separated easily from its carrier or base and information in machine-readable form can be transferred easily from one medium to another. Terms such as "original," "duplicate," and "record copy" become meaningless because electronic information can be passed from one medium to another and duplicated easily. The possibility of transferring information in machine-readable form from one type of storage medium to another also creates a wider variety of options for the cost-effective storage and preservation of information. Therefore, an important element of an expanded records management program is determining which storage formats and media are most appropriate to meet the legal, fiscal, administrative, and research needs for access to information.

The ease of manipulating information in machine-readable form has special advantages for users. Researchers who wish to examine a sample of cases, study groups or phenomenon with special characteristics, or produce statistical studies or tabulations can do so much more easily using data in machine-readable form. The U.S. National Archives, the National Archives of Canada, and many universities have developed data archives programs to meet researchers' requests for machine-readable records. Not only do researchers prefer records in machine-readable form for certain types of research, but some important historical records are now only created in computerized form.[11]

Machine-readable records will require careful handling because of their unique physical characteristics. Currently, magnetic tape is the most common medium for the long-term retention of data in machine-readable form. Unlike paper or microfilm, this medium is erasable and reusable. Magnetic media are less durable than paper or microfilm. Therefore, machine-readable records require special storage, careful handling, and regular maintenance. Computer

[11] For description of these programs, see Harold A. Naugler, "The Machine-Readable Archives Program at the Public Archives of Canada: The First Five Years," and Charles M. Dollar, "Machine-Readable Records of the Federal Government and the National Archives," in *Archivists and Machine-Readable Records*, Carolyn Geda, Erik Austin, and Francis Blouin, Jr., eds. (Chicago: Society of American Archivists, 1980); and Carolyn Geda, "Social Science Data Archives," *American Archivist* 42 (April 1979), pp. 158-66.

hardware and software are essential for retrieval of information from EDP records and for interpretation of their contents. When records are retained in machine-readable form, records custodians must ensure that the fragile medium is maintained properly and that the means to access the information including hardware, software, and adequate documentation are available. Without all of these elements — the information and the means to access it — agencies will not be able to meet records retention requirements.

### 1.3 THE ORGANIZATIONAL CONTEXT

The organizational context for the creation, storage and use of machine-readable records also differs from that for traditional hard copy documents. With large mainframe-based systems, responsibility for the systematic analysis of information needs, the design of automated systems, the processing and storage of data, and its ultimate retention or disposal usually is divided between users and data processing units. Users define the informational content of the records and may specify the formats for output reports. EDP units usually determine the physical format of machine-readable records and control their storage.

With micro-computer based systems, end users decide how to organize and access data in the system, what documentation (if any) they will create to facilitate retrieval of information by outside users, and when they will delete data from the system. Unless data management and retention guidelines are in place, every user becomes an information manager, deciding how to set up their electronic filing systems, what information to store there, and how long to keep it. Retention requirements for machine-readable records should be determined by users, but users often lack the technical knowledge and familiarity with machine-readable records to make sound retention recommendations. Therefore, EDP units often suggest data retention guidelines and usually have primary responsibility for implementing them. Because responsibility for the creation, storage, management and preservation of machine-readable records is shared, users, data processing units, agency records managers, and SARA staff must develop reasonable ways to divide these responsibilities and to coordinate information management activities.

PART TWO:

# OBJECTIVES FOR IMPROVED MANAGEMENT AND PRESERVATION OF EDP RECORDS

This section of the report discusses seven basic objectives for improved management and selective archival preservation of EDP records during the next five years. These objectives are grouped in three general categories which parallel the three Bureaus in SARA's State Government Records Division: Records Analysis and Disposition, Agency Services, and Archival Services. Because information management issues cross many constituencies, key links must be established between agency administrators, program managers and other users of EDP systems, records managers, agency information specialists, and SARA staff to achieve the objectives discussed below. Unlike a model used in some other states, no single body in New York State government can or should assume overall responsibility for the processing, storage, management and selected preservation of machine-readable records. Instead, the activities of various interagency groups and agency specialists must be coordinated at both the policy and operational levels in order to resolve the key information management issues discussed in Part One of this report.

The coordination and cooperation of several existing organizations in State government is necessary to assess existing records and information management practices and policies for EDP records and to recommend changes where needed. Currently, development and oversight of agency records management initiatives is being coordinated by the Governor's Office of Management and Productivity through the Records Management Improvement Project. As agencies inventory their records, develop records management plans, and implement records management programs, it is essential that records in machine-readable form are included. The New York State Forum for Information Resource Management, representatives of the data processing units such as the Interagency Committee on EDP, and SARA staff should work together closely to ensure that agencies develop provisions for improved management of EDP records. Closer involvement of EDP Directors in the Governor's Records Management Improvement Project would provide insights into the problems that EDP units face in managing the records they process.

Over the next five years, SARA will work closely with the Governor's Office of Management and Productivity, the Forum for Information Resources Management, and advisory groups of agency records managers and data processing staff to improve records management programs in New York State and to ensure that machine-readable records are included in these programs. To assist agencies in meeting the challenges of new information technology, SARA will promote the development of guidelines and procedures for management of machine-readable records, provide training and technical assistance, assist in evaluating new record keeping technologies, and recommend improvements in centralized storage facilities where needed. These efforts will build on the work of the Governor's Records Management Improvement Project. SARA will also develop its capacity to identify, select, accession, and preserve historically valuable EDP records at the State Archives.

Within agencies, greater effort must be made to improve records management and additional resources must be devoted to records management activities for all formats of information. SARA can provide coordination, training and technical assistance, but these centralized services will be effective only if they are complemented by increased involvement and support from agencies for records management activities. In the area of machine-readable records, EDP units must play a key role in helping agency program offices and records managers to develop retention guidelines and schedules. Furthermore, EDP Directors must be committed to continuing involvement in agency records and information management programs in order to ensure that retention guidelines are implemented as specified.

### 2.1 RECORDS ANALYSIS AND DISPOSITION

Improvements in the management of EDP records will require increased centralized services for records management to assist agencies with analysis of their records. In the area of records analysis and disposition, SARA will develop guidelines for the retention and disposition of EDP records, integrate these guidelines into revised records disposition authorization procedures, assist agencies in the development of appropriate retention periods of EDP records, and appraise potentially archival records in machine-readable form. Overall responsibility for implementing EDP records management programs will remain with State agencies. Therefore, many of the objectives discussed in this portion of the report have organizational, management,and resource implications for agencies. This report assumes that SARA will be unable to meet many of its records management and archival responsibilities without improvements in agency records management programs.

**OBJECTIVE 1: Develop guidelines and assist Executive Branch agencies to implement programs which identify, describe, and schedule the disposition of data from automated information systems.**

Traditional records management concepts and procedures provide a useful starting point for the identification, description, and retention scheduling of EDP records. As was discussed in Part One, information recorded in digitized form can be produced on paper or microfilm, and hard copy documents overlap in content with the information in data files and data bases. Therefore, the management of EDP records should be coordinated with management and retention of hard copy documents. Policies and procedures for improved records management must include both EDP records and hard copy documents in order to assist users, records managers, EDP specialists, and archivists to determine the most appropriate format(s) for retention of information.[12]

Guidelines for the identification, description, and scheduling of data from automated systems must be developed to provide the framework for integrating machine-readable records into the records disposition audit and review process. Such guidelines should help agencies to identify EDP records that can be disposed after a relatively short retention period, records that are vital, and information systems that are likely to produce records with long-term legal, research or historical value. In order to avoid undue burden on EDP units which may be required to supply descriptive and technical information about machine-readable records, it is also important that guidelines for the management of data in EDP

[12]Hard copy documents, for example, may be required if they contain signatures or other evidence of legal certifications. In other cases, retention of information on magnetic tape or some other machine-readable medium may provide a cost effective alternative to the retention of large volumes of hard copy documents.

systems use an approach that is acceptable to the EDP systems community. SARA staff will need advice from representatives of EDP units so that existing EDP systems management tools can be applied wherever they are appropriate. Section Three of this report outlines several specific activities which will lead to the gradual integration of EDP records into broader records management programs during the next five years.

An expanded records management program will rely on agencies to inventory, describe, and schedule their EDP records in conjunction with records management initiatives. In submitting records retention schedules and disposition requests, agency records management staffs will be expected to coordinate management of EDP and hard copy records. SARA archivists and records analysts must also be prepared to review comprehensive records schedules which include records in several formats and to provide advice to agencies on the retention of machine-readable records.

The reviewing bodies for disposition requests should insist that schedules submitted by agencies cover agency records in all physical formats, taking into account the need for immediate space savings through the elimination of unneeded paper records. Representatives of the Department of Law and the Office of the State Comptroller, who review disposition requests, should also consider the effects of electronic records keeping on the legal and fiscal interests of the State and its citizens. Basic policy issues, such as the admissability of machine-readable records as evidence in courts of law, need to be addressed by Department of Law staff with expertise in this area. Similarly, guidelines from the Office of the State Comptroller regarding the fiscal and audit considerations of financial records maintained in machine-readable form require clarification.

**OBJECTIVE 2: Ensure that State government programs and operations are adequately documented for legal, fiscal, administrative, and historical purposes.**

With the increased use of computers, a significant portion of agency information resources exist in machine-readable form. These machine-readable records contain vital documentation of government programs and operations. Sound information and records management practices must be applied to machine-readable records to ensure that adequate documentation of State government operations is available to meet legal, fiscal, administrative, and historical needs.

In the past, analysis of the need to keep or destroy government records generally has not occurred until a large volume of records has accumulated or until an agency requested permission to destroy records. This approach to records management is unworkable for EDP records and is increasingly impractical for managing paper records. Records and information managers agree that retention periods for machine-readable records should be established as early as possible in the record's life cycle so that records can be managed from the point at which they are created until they are destroyed or transferred to the State Archives for preservation. This is especially important for EDP records because of their fragile storage media and the need for hardware and software to access the data.

During the first two years of an expanded records management program, SARA staff in the Records Analysis and Disposition Bureau will reassess the timing of records analysis and appraisal studies. Agencies will be encouraged to develop records retention schedules for new information systems as they are developed and for existing records as early as possible in their life cycle. These retention schedules will provide basic information needed by the agencies that review disposition requests to ensure that government operations are documented adequately and to improve the economy and efficiency of the information storage and retrieval. Retention

schedules may also form the basis of a central inventory of data which is needed to enhance data sharing and reduce data redundancy among agencies. The Records Analysis and Disposition Bureau will also undertake a major investigation of the impact of office automation on record keeping and recommend appropriate policies and guidelines for managing records from decentralized systems where needed.

### 2.2 AGENCY SERVICES

In order to meet their responsibilities for better management of EDP records, agency staff need training, technical assistance, and adequate off-site storage facilities for magnetic media. In the area of agency services, SARA will provide records management training and technical assistance to agency records managers and EDP specialists, evaluate the role of the State Records Center and other facilities in providing off-site storage for machine-readable records, and monitor changes in record keeping technology.

**OBJECTIVE 3: Increase awareness—among records managers, program managers, EDP specialists, and users—of the primary and secondary values of data, and of policies regarding data access, retention, and disposition, and long-term preservation.**

One element of an expanded records management program will be training and technical assistance for agency staff in records management. EDP professionals generally are unfamiliar with the records management and archival techniques used to determine the appropriate retention period for records and to develop retention and disposition schedules. Therefore, training programs should address the special needs of users and EDP specialists by teaching them how to integrate records retention scheduling into systems design methodologies and how to identify machine-readable records which are likely to have long-term value for internal agency administration or for outside research use.

Records officers and program managers will also need training in the identification, description and scheduling of EDP records. Even though EDP systems specialists will provide some of the descriptive and technical information about electronic records, records managers and program managers will have to analyze how these records are used in their agencies. Records officers will also provide a key coordinating role in linking EDP records to related hard copy documents. Moreover, with the increased use of office automation and other decentralized technologies, program officers and records managers will have to assume greater responsibility for managing records produced in automated information systems regardless of their physical format.

**OBJECTIVE 4: Enhance the ability of State agencies and SARA to manage and, where appropriate, to preserve records created with modern information technology.**

With frequent changes in computers and information technologies, there is also a need for ongoing technical evaluation and assistance to help agencies evaluate the records management and archival implications of new technology. Currently, for example, agencies need advice and assistance with the information management aspects of office automation and optical disk technology. The impact of stand-alone and network systems on the creation, storage and retention of basic office communications needs to be further evaluated. Records and information management guidelines for the electronic and hard copy products of office automation systems should be developed to help agencies ensure that their policy formation and decision-making processes are adequately documented.

Several agencies are investigating the use of optical disks as an alternative to hard copy or microfilm storage. Feasibility studies investigating the use of this technology must address not only the immediate cost implications of the use of optical disks, but also the long-term records management and archival implications. In order to keep pace with such changes in information technology, SARA will monitor developments in informa-

tion technology and storage media and provide advice and technical assistance to agencies contemplating the use of new storage techniques.

**OBJECTIVE 5: Evaluate the role of the State Records Center in data storage, security, and preservation, and provide improved services where indicated.**

The Records Center provides off-site storage for back-up files and inactive magnetic tapes. During the first two years covered by this plan, SARA should assess the adequacy of Records Center tape storage facilities and services. Because magnetic tape is a fragile medium which must be stored in areas with stringent temperature and humidity controls, conditions in the Records Center tape storage vaults should be evaluated. Security systems should also be evaluated to ensure that vital data is secure and that the possibility of unauthorized access to restricted data is minimized.

Even under optimum storage conditions, magnetic tape deteriorates after a period of 2 to 10 years. Regular maintenance procedures, including rewinding of inactive tapes, checking for tape deterioration, and the periodic transfer of data to new tape, are essential to ensure the readability of data when its retention exceeds five years. Currently, no centralized service is available to provide regular maintenance for magnetic tapes or to monitor their readability over extended retention periods. There may be a role for expanded Records Center services to provide such maintenance services for inactive machine-readable records.

SARA should assess the need for centralized maintenance and conservation services for EDP records at the Records Center. Improvements may be needed to ensure that temperature and humidity in tape vaults meet recommended standards for long-term retention of magnetic tape. Procedures should also be developed to ensure that magnetic tapes stored at the Records Center for purposes other than short-term back-up protection are covered by Records Disposition Authorizations as is currently required for hard copy records.

### 2.3 ARCHIVAL SERVICES

The long-term preservation of selected, valuable machine-readable records will require expanded archival services. Objectives for improved archival services relate primarily to the State Archives' responsibility to select, acquire, preserve and make available State government records of enduring value. This strategic plan is designed to phase in these services gradually over the next five years. SARA will begin by acquiring and describing approximately 100 valuable machine-readable data files already appraised by the Special Media Records Project staff. During the next five years, SARA will develop the capacity to accession, describe, and prepare for storage data files with enduring value, locate adequate storage space, develop a tape monitoring and maintenance program, and provide reference services for its machine-readable records holdings.

**OBJECTIVE 6: Develop the capacity of the State Archives to accession, describe, and make available for research use machine-readable records holdings.**

In order to ensure that selected, valuable records are preserved for future research, the State Archives must acquire the resources and develop the capacity to process and describe archival records in machine-readable form. It is absolutely essential to obtain access to data processing services during the first year of this plan in order to begin accessioning more than 100 data files which have already been identified as archival. In order to develop the capability to acquire valuable EDP records on an ongoing basis, the Bureau of Archival Services, Collections Management Unit, will develop technical specifications and procedures for the transfer of data files to archival storage from State agencies. Such guidelines will indicate the formats that agencies must use to transfer data to the Archives and set minimum standards for docu-

mentation that must accompany archival data files.

Based on the anticipated volume of accessions, the Collections Management Unit will also establish processing guidelines for archival data files. In the past, data archives have followed extensive verification procedures for all accessions. These procedures involve comparing print-outs of selected records against codebooks to determine the accuracy of the data and documentation, performing some simple statistical analyses to check for consistency of coding schemes, and taking corrective measures where indicated. Some data archives are reassessing the feasibility of performing such a thorough analysis for all accessions.[13] In establishing accessioning guidelines, the Collections Management Unit will consider whether different levels of processing are appropriate depending on the value and uniqueness of the data, the costs of verification and possible remedial actions, and anticipated research uses. During the planning period, the Collections Management Unit will also establish descriptive standards which will meet internal Archives' administrative needs to control archival data files and will conform to national standards for the exchange of descriptive information about archival holdings.

During the next five years, the Archives anticipates a gradual increase in reference requests for machine-readable records. The number of data files acquired initially will be modest and researchers may be unaware that the State Archives preserves data in machine-readable form. Therefore, special efforts must be made to publicize the State Archives' data holdings. In order to encourage research use of the Archives' machine-readable record holdings, the Bureau of Archival Services, Reference Services Unit, in conjunction with the SARA External Programs Division, will submit notices and articles to the appropriate journals, and produce descriptive brochures, guides and other finding aids where appropriate. The Collections Management Unit will also enter descriptive information about SARA machine-readable records holdings into the appropriate national databases as the records are acquired and described. During the next five years, the Reference Services Unit will also develop the capacity to meet researchers' requests as they occur. This will involve establishing procedures for copying tapes in conjunction with a data processing service bureau and gaining familiarity with descriptive materials for machine-readable records. Policies and procedures will also be developed to balance rights of access with the protection of confidential data, especially data which may affect personal privacy.

**OBJECTIVE 7: Develop the capacity of the State Archives to store machine-readable records and perform maintenance and preservation activities for EDP records.**

Preservation activities for machine-readable records during the next five years will be limited primarily to identifying and establishing appropriate storage facilities and developing a tape preservation monitoring system. Locating adequate storage space for master and back-up copies of archival tapes is a top priority because the Archives already has acquired a few data files and anticipates additional accessions in the coming year. The storage of magnetic tape in substandard storage areas increases the risk of data loss and will add to preservation costs in the future. Ideally, master tapes will be stored in a tape library associated with the data processing service bureau that the Archives selects. If the service bureau cannot provide secure and environmentally controlled storage, an alternative site will be required. This may require installation of environmentally controlled vaults at the State Archives. In addition, an off-site storage lo-

[13]Public Archives of Canada, "Five Year Plan for the Managing and Archiving of Machine Readable Data in the Public Archives of Canada," Ottawa, 1986, pp. 56-57.

cation must be identified for back-up copies of tapes. The tape storage facilities at the Records Center will be evaluated for this purpose. Long-range plans for improved State Archives facilities must accommodate the storage requirements of special media records.

Because the State Archives is just beginning to acquire machine-readable records, only basic conservation activities are anticipated during the next five years. Tape maintenance services, such as periodic rewinding of tape, checking for errors, and transfer of data to new tape will be provided by a data processing service bureau. The Archives' Preservation Unit must develop a tracking system to monitor tapes and to ensure that maintenance procedures are performed on a regularly scheduled basis. Such a system should be in place by the second year of the planning period to accommodate current and anticipated accessions. The Preservation Unit will also assume overall responsibility for investigating alternative storage media to magnetic tape. Although no cost-effective alternative currently exists, developments in optical disk technology and other storage media should be monitored as potential alternatives to magnetic tape storage.

PART THREE:

# SPECIFIC ACTIVITIES AND RESOURCE REQUIREMENTS OF THE FIVE-YEAR PLAN

Part Three of this report outlines 27 specific activities that must be undertaken during the next five years (FY 1988-89 to FY 1992-93) to integrate machine-readable records into State government records management programs and to develop the capacity at the State Archives to select, acquire and preserve records with enduring value in machine-readable form. The activities establish a workplan to accomplish the objectives described in Part Two. As with the objectives, the activities are grouped into three categories which correspond to the three Bureaus in SARA's State Government Records Division: Records Analysis and Disposition, Agency Services, and Archival Services. For each activity there is a time frame and an explanation of why each issue should be addressed and how each activity should be undertaken. Although each activity is assigned to a specific SARA unit, many activities will require assistance from State agencies and additional resources in agencies to improve their own management of EDP records.

**Summary of Objectives and Activities**

RECORDS ANALYSIS AND DISPOSITION: This Bureau guides the adequate documentation of New York State government, provides analysis and review of all Executive Branch records disposition, coordinates approval of records disposition schedules and requests with the Department of Law and the Office of the State Comptroller, and conducts archival appraisal of the records of all three branches of government. This plan sets two objectives for this Bureau: 1) to develop guidelines for the identification, description and retention scheduling of EDP records so that machine-readable records are included in the records retention and disposition review and approval process, and 2) to ensure that agency programs are adequately documented for legal, fiscal, administrative, and historical purposes. Specific activities to accomplish these objectives include the development of guidelines and procedures for scheduling EDP records, completion of a survey of centralized EDP systems, development of guidelines for identifying potential archival records in electronic formats, a study of the implications of office automation systems for records keeping and records retention, and an evaluation of access policies for records in machine-readable form. In addition, this Bureau will analyze and appraise EDP records as early as possible in their life cycle and conduct special appraisal studies of older and endangered machine-readable records. The goal of this plan is to develop basic policies, guidelines and procedures by 1990 for EDP records produced by centralized automated systems and to have policies and guidelines in place for office automation systems by 1992.

AGENCY SERVICES: This Bureau operates the State Records Center for the storage of inactive records, provides other centralized services such as micrographics services, develops education and training programs in all aspects of records management, develops standards, and provides technical assistance in the use of technologies and in sound records management techniques. The objective of this plan is to ensure that SARA services in these areas meet the needs of agencies for electronic records management. Specific activities will include the development of educa-

tion and training materials on data management and retention scheduling of EDP records, the dissemination of guidelines on tape maintenance and preservation, and the evaluation of standards for tape storage and data interchange. This Bureau will also monitor new record keeping technologies, such as optical disk, and provide assistance to agencies in use of these technologies. To ensure that off-site storage for back-up tapes and inactive data files meet the special needs for maintenance and storage of EDP records, this Bureau will assess current conditions. The assessment may recommend improvements in storage vaults and additional tape maintenance and preservation services.

ARCHIVAL SERVICES: This Bureau serves as New York State government's archival program. It accessions, arranges and describes, preserves and provides reference services for the permanently valuable records of New York State. Because New York State's permanently valuable records now include records in machine-readable form, the objective of this plan is to ensure that the State Archives develops the capacity to acquire, preserve and make EDP records available. During the first two years of the plan, the Bureau's Collections Management Unit should secure computer services for accessioning and processing EDP records, develop descriptive standards for these records, and obtain sufficient staff to handle accessioning and descriptive work. As the State Archives acquires permanently valuable records in machine-readable form, the Bureau's Reference Services Unit should publicize their availability to researchers and develop the capacity to meet research requests. The Preservation Unit should ensure that data processing services are available to perform basic tape maintenance and preservation activities, and should investigate alternative storage media for electronic records. By the end of the planning period, the State Archives will need a carefully controlled storage area to house 150 to 200 reels of tape and to accommodate a collection of several thousand tapes in the future. Long-range facilities planning must include provisions for a much larger volume of material, based on estimates that are revised as the Archives acquires more records in electronic formats.

After the description of objectives and activities for each program area, an estimate of resource needs is provided. Estimates of resource needs at SARA include personal services, supplies, equipment, and contract services. Resource estimates are also provided in tabular form at the end of the report. The resources needed for some activities cannot be estimated at this time because more detailed studies or needs assessments must first be completed. Moreover, this plan does not attempt to estimate the resources that agencies will need to improve management of their EDP records or to assist SARA staff with several of the activities outlined in the report. SARA will need assistance and advice from the NYS Forum for Information Resource Management, the Interagency Committee on EDP, the Governor's Office of Management and Productivity, and from other advisory bodies to ensure that policies and programs for managing electronic records meet the needs of agencies and of State government. Once overall policies and guidelines are in place, agencies must devote additional resources to the analysis and retention scheduling of EDP records, to assessments of internal policies, and to staff training. Many agencies may need additional staff in both the EDP area and in records management to develop retention and disposition schedules and to better manage and preserve electronic records. SARA will need agency assistance to identify EDP records with archival value, to compile adequate documentation, and to transfer permanently valuable EDP records to the State Archives for preservation.

# Description of Specific Activities

Program Area: RECORDS ANALYSIS AND DISPOSITION

Objective 1: Develop guidelines and assist all Executive Branch agencies to implement programs which identify, describe, and schedule the disposition of data from automated information systems

*Activity 1:* *Support and promote development of government-wide policies and guidelines for retention and disposition of records from automated information systems.*

Time Frame:

| | |
|---|---|
| 1988-89 to 1989-90 | Policy Development |
| 1990-91 to 1992-93 | Implementation |

Explanation: During the past few years, New York State has recognized problems with the inadequacy of its records management programs. Several coordinating bodies have been established and new initiatives are being undertaken to address these deficiencies. All Executive Branch agencies are participating in the Records Management Improvement Project sponsored by the Governor's Office of Management and Productivity. Recently, managers from many State agencies have joined together to form the New York State Forum for Information Resource Management. The Forum promotes policies and practices for effective use and management of information resources in State government. In order to improve coordination between records management and archival activities, the State's records management and archival functions were consolidated with the transfer of statewide records management program responsibilities to the State Education Department.

Efforts to improve records management and to establish effective information resource management programs should be comprehensive. Guidelines, policies, and procedures for records management programs must include EDP records because an increasing portion of the State's public records exist in electronic formats. Information resource management programs must consider broad policy issues of access, use, dissemination, maintenance, and preservation of information in both electronic and hard copy formats. SARA should work with the Governor's Office of Management and Productivity, the NYS Forum for Information Resource Management, and the Interagency Committee on EDP to ensure that these efforts are comprehensive and coordinated. These groups should recommend government-wide policies where needed to improve management and archival preservation of records from automated information systems.

*Activity 2:* *Develop guidelines for records retention and disposition that integrate EDP records into a revised review process under subdivision 11, Sec. 57.05 of the Arts and Cultural Affairs Law.*

Time Frame: 1988-89 to 1990-91 Development
1991-92 to 1992-93 Implementation

Explanation: During the past two years, the Special Media Records Project has acquired some experience in the scheduling of data from automated systems. The project staff recommend that agencies approach the development of retention schedules for machine-readable records from a total information system perspective whenever possible. Agency EDP staff, program staff, and records officers should analyze their automated information systems to determine how long specific files, data bases, or data elements need to be retained, which formats are most appropriate for retention, how information will be maintained and made accessible during its retention period, and which records, if any, should be transferred to the State Archives for preservation. Disposition analysis should consider the retention of machine-readable and hard copy records together. Whenever possible, decisions regarding the retention, maintenance and disposition of data should be made during the design phase for new automated systems.

While these general principles can aid the development of policies for scheduling records from automated systems, agencies need more specific guidelines for determining appropriate retention periods and identifying EDP records with potential archival data. Staff in the Bureau of Records Analysis and Disposition should draft guidelines for retention and preservation of EDP records. These guidelines might include a General Retention and Disposition Schedule for data processing records that are common to all agencies and have short-term retention requirements. Draft guidelines should be reviewed by agency records officers, program managers, and EDP staff, and revised to address organizational, technical, and practical problems in the agencies.

Once guidelines are established, agencies should develop internal mechanisms to ensure that EDP records are covered by adequate retention and disposition schedules. Scheduling EDP records will require additional resources in agencies for records management, systems analysis, and data management activities. These resources are not included in the resource estimates below.

An essential element of improving the retention and disposition of EDP records is to ensure that records in electronic and other machine-readable formats are incorporated into the records disposition review process. In the past, machine-readable records have not been covered by disposition schedules or approved Records Disposition Authorizations (RDAs) as required by Subdivision 11, Sec. 57.05 of the Arts and Cultural Affairs Law (formerly Sec. 186 of the State Finance Law). To remedy this situation, the parties which review Records Disposition Requests should insist that agencies develop retention schedules for their EDP records.

### *Activity 3:* *Conduct a survey of centralized automated information systems in all Executive Branch agencies.*

Time Frame: 1988-89

Explanation: The Special Media Records Project compiled an inventory of all automated information systems supported by the central EDP units in 19 Executive Branch agencies. This information provided Archives staff with an overview of the major data systems in each of the agencies and formed the basis for identifying priority areas for follow-up analysis and appraisal.

During the current year, all agencies will complete an inventory of their records as part of the Governor's Records Management Improvement Project. The survey form, (OMP-1) asks each program unit to identify the data processing system or application, if any, used to create records that are stored in program units.

The inventory did not collect information about the automated systems supported by central EDP Units. In order to fill in this gap, the SARA Bureau of Records Analysis and Disposition should sponsor a survey of centralized EDP systems (except in those agencies already surveyed by the Special Media Records Project). This inventory information, combined with information already collected for the Records Management Improvement Project, will allow agency records officers and SARA staff to identify centralized automated information systems and link hard copy input and output records to a system or application. The inventory information would also support the development of a State government data inventory currently under consideration by the New York State Forum's Information Clearinghouse Working Group.

*Activity 4:* ***Develop and implement guidelines for the management, scheduling, and disposition of data in office automation environments.***

Time Frame:

| | |
|---|---|
| 1988-89 to 1989-90 | Analysis |
| 1990-91 to 1992-93 | Program Development and Implementation |

Explanation:

In the past few years, records managers and archivists have begun to study the impact of office automation systems (stand-alone word processors and document-based communications networks) on the management, retention, and preservation of records. Pilot studies indicate some significant differences between office automation environments and centralized data processing systems.[14] Because office automation systems are generally user-driven, responsibility for the retention and disposition of the records they create often rests with end users.

The State Archives and Records Administration has not conducted research on the management or preservation of records from automated office systems in New York State agencies. Little is known about the extent to which office automation systems are being used, the types of information created and stored in office automation environments, or the current policies, if any, regarding the management and disposition of records from these systems.

During the first two years of the planning period, SARA will undertake a study of information management and records disposition practices in program units with major office automation installations. Such a study will determine whether documents that formerly were created only in hard copy are now produced in machine-readable form; assess changes in information management and disposition practices; analyze the impact of data transmission capabilities on the revision, distribution and retention of documents; and evaluate the indexing and retrieval capabilities of the system. Selected data from the Records Management Improvement Project inventory will be used to assess the extent of office automation in Executive Branch agencies.

Based on this study, SARA staff will work with agency program managers and records managers to draft guidelines for management and retention of records created in office automation environments. Such guidelines will define various stages in the production of correspondence and policy documents; help agencies determine when drafts as well as final reports need to be retained in order to document the formation of policy; and recommend appropriate formats (machine-readable, microform, hard copy) for long-term retention of information from office automation systems. The proposed guidelines will be tested in a pilot agency.

[14]See especially the interim Reports of the Public Archives of Canada/Department of Communications Information Management Working Group, *Interim Reports*, prepared by John MacDonald, chairman, Ottawa, Public Archives of Canada, June 1984 and September 1985.

Based on the experience of this pilot project, general guidelines will be developed for implementation in all Executive Branch agencies. Such guidelines will help agencies identify when disposition requests must be filed for the machine-readable as well as hard copy documents, and provide policies and procedures for implementing the disposition guidelines.

**Objective 2:** **Ensure that State government programs and operations are adequately documented for legal, fiscal, administrative, and historical purposes.**

***Activity 5:*** ***Assess the timing and manner in which records analysis and archival appraisals are conducted in order to encourage comprehensive retention scheduling as soon as possible after records are created.***

Time Frame: 1988-89 to 1989-90

Explanation: Improved records management practices for New York State government agencies should be based on sound information management principles. Records and information managers recognize that, given the impermanence of the storage media on which data is stored, hardware and software dependency issues, and the relationship between machine-readable and hard copy records, the retention periods for all records created by automated systems should be established as early as possible in the record's life cycle.

During FY 1988-89 and 1989-90, staff of the SARA Bureau of Records Analysis and Disposition will reevaluate the timing of records analysis and appraisal studies. This analysis will cover both EDP and hard copy records and assess the practical problems of evaluating records as soon as possible after they are created. SARA staff will develop procedures to ensure a more timely analysis of new records systems as they are designed, including mechanisms for learning about new information systems before they are implemented.

***Activity 6:*** ***Develop guidelines for the identification and appraisal of EDP records with potential archival value and identify priority appraisal projects.***

Time Frame: 1988-89 to 1989-90

Explanation: The Special Media Project staff gained considerable experience in evaluating the archival value of EDP records. The Project's Preliminary Report estimated that as many as 35 percent of the major, centralized EDP systems may produce some data with archival value. The Report also suggested that survey files, registration systems, case management systems, license-permit systems, and some general operational systems were most likely to produce data of archival value because they contain detailed social and demographic information about individuals affected by State programs; information about important economic, social and scientific phenomenon; and information that documents the policies and programs of State government.

Five major EDP systems were analyzed in detail by Project staff to determine precisely which data files, if any, contained information of sufficient historical or other research value to warrant long-term preservation in machine-readable form. [15]

[15]The systems studied were the State Education Department's Basic Educational Data System, the New York State Economic Database in the Department of Commerce, the Computerized Criminal History System (CCH/OBTS) in the Division of Criminal Justice Services, the Vital Records System in the Department of Health, and the Division of Equalization and Assessment's Private Utility and Industrial Property Valuation System. Copies of these studies are available on request from the State Archives and Records Administration.

Thorough analysis of the systems enabled the Project staff to better estimate the amount of appraisal work needed to identify and preserve value archival records in machine-readable form. After appraising five systems, the Project staff concluded only three produced data which should be preserved by the State Archives. This suggests that between 15 and 20 percent of centralized EDP systems might produce archival records, rather than 35 percent as originally estimated.

Given the extensive use of computer technology by State government agencies, it is important to stress that this is still a very large number of systems. Possibly as many as 150 to 200 centralized EDP systems produce some data with archival value. Therefore, guidelines are needed to identify those systems that are most likely to produce valuable data for future research.

Based on the Special Media Project studies, it appears that data files created and used by agencies for their own internal research purposes are the most likely to have future research applications. This includes survey files and special research files which are often extracted from larger database for research and policy evaluation by agencies.

Even with additional appraisal staff, only a small percentage of the EDP systems which may produce archival records can be appraised during the next five years. Therefore, during the first year of this plan, a list of top priority systems for appraisal will be identified. In establishing priorities, SARA staff will take into account the uniqueness of the data, its age, the extent to which it may be endangered in its current environment, and the likelihood that a detailed study will yield either valuable archival data files, major records management improvements, or both. This list will be used to allocate staff for special appraisal projects throughout the planning period and may provide the basis for a special project to preserve selected, valuable EDP records. Agencies need additional guidance to help Archives staff identify other criteria which might be used to identify machine-readable records with potential archival value. SARA staff should develop a set of general guidelines for identifying EDP records with potential long-term value. This information will be provided to agencies to assist them with scheduling their EDP records. General guidelines for identifying EDP records with long-term value, however, will not eliminate the need for thorough analyses of disposition schedules for major EDP systems by staff of the Bureau of Records Analysis and Disposition.

As agencies begin to submit retention and disposition schedules for EDP records, it is likely that a significant number of valuable, endangered machine-readable records will need immediate appraisal and timely transfer to the State Archives for preservation. Based on the results of initial inventory and records scheduling activities in agencies, a special initiative may be needed at SARA, beginning in 1990-91, to identify and preserve selected, valuable EDP records that are in danger of loss or deterioration in State agencies.

*Activity 7:* ***Conduct special appraisal studies of several major information systems each year.***

Time Frame: Ongoing

Explanation: To accomplish the State Archives' mission to identify and preserve records of lasting value for historical and other research, the Bureau of Records Analysis and Disposition must develop the capacity to appraise machine-readable records from several major automated systems each year. Specific systems to be appraised will be selected from the list of priority appraisal projects (see Activity 6) and based on agency records management improvement plans and schedules submitted by agencies.

*Activity 8:* ***Include machine-readable records in special initiatives to improve records management and archival programs in the Legislative and Judicial branches of government.***

Time Frame: Ongoing

Explanation: The Legislative and Judicial branches of government face records management problems for both traditional files and machine-readable records that are similar to those encountered in Executive Branch agencies. The Senate, the Assembly, and the Unified Court System rely heavily on automated information systems for records creation, retrieval, and storage. However, no comprehensive guidelines exist to promote records management or ensure archival preservation of valuable records in these branches.

Two initiatives are under way to address records management and archival problems in the Legislature and the Courts. Since 1984, the State Archives has worked with the Office of Court Administration on a Judicial Records Project which has resulted in records retention schedules for all of the State's major types of courts. During the next five years, as these schedules are tested and implemented in court sites, OCA records managers will likely encounter and need to address issues and problems with EDP records which parallel those of the Executive Branch.

A preliminary assessment of archival and related records management needs in the Legislature, begun by SARA in 1987, indicates that the use of computers is widespread. In the Assembly, for example, all assemblymen have personal computers for producing correspondence. Offices, such as the Task Force on Demographic Research and the Legislative Bill Drafting Commission, appear to produce important records in machine-readable form. In order to ensure that the full range of records management and archival issues is addressed in the Legislature, machine-readable records must be included in SARA's plans to identify and preserve valuable legislative records.

## RECORDS ANALYSIS AND DISPOSITION: SUMMARY OF RESOURCE NEEDS

**During the first two years, the Bureau Chief will devote 15 to 20 percent time to coordinating and supervising the development of data retention guidelines, and to revising the records disposition review and approval process to accommodate EDP records. In addition, the Bureau Chief will consult with the Head of the Collections Management Unit in the Bureau of Archival Services on obtaining adequate data processing services and on developing accessioning and descriptive procedures for archival machine-readable records.**

**One additional Associate Archivist is needed in this Bureau to develop guidelines for identification and appraisal of EDP records with potential archival value and to conduct appraisal studies of valuable and endangered EDP records. This Archivist will also assist the Collections Management Unit with transfer to the Archives of more than 100 data files that were identified as archival by the Special Media Project. During FY 1988-89 SARA should fill a vacant temporary services position for an Associate Archivist. This position should be transferred to permanent State funding beginning in FY 1989-90. Approximately 25 to 30 percent of the time of an Associate Records Analyst is also needed to develop guidelines for retention and disposition of EDP records and to conduct a survey of centralized automated information systems. The Associate Archivist for Legislative Records will spend approximately 10 percent time evaluating the records keeping and archival implications of EDP records in the Legislature. Beginning in 1990, this Bureau will need one additional Associate Analyst to develop guidelines for records produced by office**

automation systems. Furthermore, all professional staff in the Bureau will review, analyze, and appraise records schedules and disposition requests from agencies that produce EDP records. Additional resources may be needed beginning in 1990 to appraise valuable data that is in danger of loss or destruction.

Program Area: AGENCY SERVICES

Objective 3: Increase awareness among records managers, program managers, EDP specialists and users of the primary and secondary values of EDP records, and of policies regarding access, retention, disposition, and long-term preservation of data.

*Activity 9:* ***Evaluate government-wide policies regarding access to records to ensure that data in automated information systems is available and that confidential data is protected.***

Time Frame:

| | |
|---|---|
| 1988-89 to 1989-90 | Development |
| 1990-91 to 1992-93 | Dissemination |

Explanation:

Access to public records is currently regulated by the Freedom of Information Law (Article 6, Public Officers Law), the Personal Privacy Protection Law (Article 6A, Public Officers Law) by specific statutes and administrative rules, and, in some cases, by federal laws and regulations. These regulations apply to all public records regardless of physical format or recording medium. More precise guidelines may be needed to provide access to records in machine-readable form. Special procedures may also be needed to protect confidential data because such information is easily duplicated and unauthorized access to computerized data banks may be difficult to trace. The complex flow of information among State agencies and between the State and federal level also raises jurisdictional issues.

A cooperative assessment of existing access, confidentiality, and security regulations and procedures for machine-readable records is needed to determine 1) whether existing policies and procedures are adequate to protect confidential information and to provide access to information that is not restricted; 2) whether access and privacy provisions are consistent among agencies and between the State and federal level; 3) whether current policies and procedures provide sufficient guidance to agency staff who must administer requests from the public to access data in automated information systems; 4) what provisions, if any, should be made regarding the transfer of machine-readable records that contain confidential data to the State Archives; and 5) what provision should be made to provide researchers with restricted access to confidential data.

Broad questions such as the definition of a "record" in a database environment and the status of software developed in-house by State agencies must be addressed. An assessment of security provisions should include the impact of increasing decentralization of data processing on the physical security and privacy protection of data.

Once current policies have been assessed and gaps and inconsistencies identified, an interagency working group, including staff from SARA and from the Committee on Open Government, should develop remedies. These may include changes in the existing law, further development and monitoring of existing procedures, more precise guidelines for agency staff, and other activities where appropriate. Once consistent policies are in place, this information should be disseminated to program officers, records managers, and EDP units through SARA and agency training programs.

*Activity 10:* ***Develop and disseminate training materials on the management, preservation, and uses of EDP records to records managers, program managers, and EDP units.***

Time Frame: Ongoing

Explanation: To expand and improve the State's records management program, SARA will provide training to records officers and program managers. In order to integrate management of EDP records into this program, records officers need training in the management and scheduling of machine-readable records. Special initiatives are also needed to provide training in the management of EDP records to data processing staff.

In fiscal year 1988-89, SARA will assess training needs and develop training proposals for all types of records management activities. To ensure that SARA educational programs adequately cover training needs for both EDP and paper records, the SARA Training and Education Coordinator should develop workshops and other training materials addressing the issues and techniques of EDP records management. Training in data management and EDP records scheduling should be disseminated to agencies as an ongoing element of the SARA training and education program.

SARA should also use consultants and contract services for state-of-the-art presentations on new issues in records and information management. Likely areas for training in the later years of the plan include electronic mail, optical disk technology, and managing information in office networks.

## Objective 4: Enhance the ability of State agencies and SARA to manage and preserve records created with modern information technology.

### *Activity 11: Integrate guidelines for scheduling EDP records into information systems design methodologies.*

Time Frame:
1989-90 to 1991-92 Development
1992-93 Implementation

Explanation: Archivists, records managers, and other information specialists agree that provisions for the management and long-term preservation of EDP records should be an integral part of systems design methodologies. Just as current automated systems planning includes specifications for the appropriate hardware and software to meet immediate information needs, systems planning should also include a retention plan for the products of automated information systems. Using this approach, agencies can develop comprehensive records retention plans for all of the records associated with automated systems. Automated systems can be designed to handle the disposition of information more or less automatically by deleting data when no longer needed for legal, fiscal, administrative or research purposes and by creating those records with archival value in a format appropriate for transfer to the State Archives (See Activity 18). Retention plans, once approved as required by law, would also bring records created and stored in machine-readable form into compliance with legal requirements for approval prior to the destruction of public records.

Beginning in FY 1989-90, the SARA Records Management Specialist for Technology and Standards will work with an advisory group of EDP specialists and agency records managers to develop guidelines to integrate the retention scheduling of EDP records into information systems design procedures. After review and testing by agencies, these guidelines should be applied to new automated information systems as they are designed.

### *Activity 12: Develop policy guidelines and procedures for preservation of data by agencies over extended periods of time.*

Time Frame: 1989-90 to 1991-92

Explanation: Some machine-readable records do not have permanent value, but still need to

be retained for extended periods of time to meet agency administrative, fiscal, and/or legal requirements. Machine-readable media, however, are subject to deterioration after a relatively short period of time. When agencies need to retain machine-readable records for more than five years, special storage and maintenance procedures should be followed in order to increase the longevity of data files and ensure that data is available throughout the specified retention period.

The SARA Specialist for Technology and Standards, in conjunction with selected agency EDP units, should develop a set of storage and maintenance procedures which will help agencies preserve machine-readable records to meet their internal information requirements. Such guidelines will specify proper temperature and humidity requirements for long-term storage, identify where adequate storage facilities are available, describe maintenance routines, encourage proper back-up procedures, and indicate when data stored on tapes, diskettes and other magnetic media should be copied to new media to ensure long-term access to the information. The guidelines will also address the need to ensure that software necessary to retrieve data in inactive storage is available throughout the specified retention period for the records or to reformat records so that they can be retrieved without access to special software.

### *Activity 13:* *Monitor technological developments in information systems and storage technology.*

Time Frame: Ongoing

Explanation: Innovations in information and storage technology promise to provide new and cost-effective ways to store and communicate information created and used by State government agencies. However, such changes will have a profound impact on the management and archival preservation of records. In dealing with major changes in information technology in the past, records managers and archivists have taken a reactive stance, developing solutions only after significant management and preservation problems were apparent and important documentation was lost.

In order to respond to ongoing changes in information technology and to assess their likely impact on records management and archival programs, the SARA Technology and Standards Specialist will monitor new information technology. Technology assessments will address two aspects of new information technologies: 1) their potential use for improved records creation, storage and retrieval, and 2) their possible implications for the archival preservation of selected, valuable government records. Office automation systems, electronic mail, optical disk technology, and data transmission systems are priority areas for assessment in the initial years of the planning period.

### *Activity 14:* *Promote government-wide procedures and mechanisms for data interchange between hardware and software systems by evaluating standards for the transfer of data from one computer system to another.*

Time Frame: Ongoing

Explanation: The incompatibility of computer hardware and software systems creates a variety of problems for data administrators, users, records managers, and archivists. Systems incompatibility exists on three levels: hardware, operating systems, and applications software. Hardware incompatibility is most prevalent in microcomputer-based systems and office automation networks. The use of diskettes of different sizes and disk drives which write data in various formats makes the exchange of data from one hardware system to another impossible in many cases. Incompatible operating systems are a further obstacle to data interchange

among systems. In office automation environments this also contributes to a proliferation of applications software and to the creation of data in formats that cannot be easily passed from one system to another. In mainframe-based systems, software incompatibility is especially common where hierarchical or relational database management systems are used.

Hardware and software incompatibility creates data interchange problems for current users of records. When two users wish to exchange information electronically, typically they must have access to the same or compatible hardware and software systems, or the data must be reformatted so that it can be accessed and read by the receiving system. The implications of systems incompatibility are also critical for records managers and archivists. If data are to remain accessible in machine-readable form over extended periods of time, archivists and records managers must ensure that the hardware or software necessary to access and retrieve the data will be available with the data. Most data archives have dealt with this problem by preserving data in a format that is as hardware-and software-independent as possible (generally as numeric data in a sequential file with fixed-length fields and records).

Archivists and records managers are reassessing the problems of hardware- and software-dependency. Downloading data from a database management system into flat files usually is a costly undertaking because it requires extensive analysis and programming, especially where utility packages are not readily available to download data. Moreover, reformatting data from a database environment designed to provide direct access to various data elements and records into a format where data can only be accessed sequentially can significantly reduce the usefulness of the data. Finally, the flat file format is inappropriate for many systems that store non-numeric data, such as cartographic databases, computer-assisted design systems, and text retrieval systems.

Some archivists and records managers are exploring alternative formats for the preservation of machine-readable data. One approach is to preserve the software necessary to access a database or a text file along with the data. This approach has generally been considered impractical because of the cost of maintaining software which may be used only occasionally. However, under certain circumstances, especially in the area of common micro-computer software packages, the costs of software maintenance may be justified if a large volume of records with long-term value are created in these formats.

A more promising alternative is the use of data interchange standards which are currently being developed by various national and international standards bodies. The ISO reference model, for example, a developing standard for data interchange between mainframe systems, shows promise of promoting standard interfaces which would allow exchange of data without reformatting between incompatible systems. Interchange formats have also been proposed and adopted by numerous federal government agencies for the interchange of textual documents between office automation systems. The EDP community, records managers, and archivists should evaluate the utility of these standards and investigate various means to promote their adoption and use by State government agencies.

## Objective 5: Evaluate the role of the State Records Center in providing data storage, security and preservation.

### Activity 15: *Assess the quality and capacity of storage areas for EDP records, including their utility for data security and vital records protection.*

| Time Frame: | | |
|---|---|---|
| | 1988-89 to 1989-90 | Assessment |
| | 1990-91 to 1992-93 | Make improvements as indicated |

Explanation: Currently the State Records Center provides off-site storage for back-up files and inactive magnetic tapes. In FY 1987-88 the Records Center handled approximately 150,000 reels of tape. This activity is growing at a rate of about 20 percent per year. In addition, approximately 22,500 inactive tapes are stored at the Records Center. In order to ensure that these records are stored under proper conditions, an assessment of the environmental quality of the storage areas should be undertaken. Magnetic tapes should be stored at as constant a temperature and relative humidity as possible. Current standards recommend a storage temperature of 65 degrees F with fluctuations in storage areas of no more than (+ or -) 5 degrees, and a relative humidity of 45 percent with a fluctuation of no more than (+ or -) 5 percent. Studies of tape deterioration indicate that storage at a constant temperature and humidity are more important than base temperature and humidity levels because fluctuations in temperature and humidity are most damaging to tapes. Temperature and humidity in the tape vaults should be monitored over a one-year period to ensure that the storage conditions meet basic quality standards. Security systems should also be evaluated to ensure that vital data is secure and that the possibility of unauthorized access to restricted data is minimized. An evaluation of this service should also review the capacity of the Records Center to handle a rapidly increasing amount of activity. Based on an assessment of these problems during the first year of the planning period, appropriate measures should be undertaken to remedy any problems uncovered by the evaluation.

*Activity 16:* ***Determine whether the Records Center should provide tape maintenance and/or conservation services for EDP records with retention periods in excess of five years.***

Time Frame:

| | |
|---|---|
| 1988-89 to 1989-90 | Assessment |
| 1990-91 to 1992-93 | Implementation |

Explanation: Regular maintenance procedures are needed to ensure the readability of data when magnetic tapes are stored for periods of more than five years. Tapes in inactive storage should be inspected periodically, rewound under constant tension, and transferred to new magnetic media to ensure the long-term readability of the data. Currently these activities, if conducted at all, are the responsibility of the agencies which create or produce the data. During the first two years of the planning period, a feasibility study of the need for centralized services for tape maintenance and conservation will be conducted to determine whether it would be cost-effective for SARA to provide such services on a centralized basis. Pending the outcome of the initial assessment, further development of tape cleaning, inspection, and copying services may be developed.

## AGENCY SERVICES: SUMMARY OF RESOURCE NEEDS

**During the first two years, an evaluation of tape storage conditions and an assessment of needs for centralized tape maintenance and preservation will require approximately 10 percent of the time of the Bureau Chief. The Coordinator for Training and Education should spend approximately 15 percent time throughout the planning period on needs assessment, and on the development and procurement of training materials for EDP records management. In addition, approximately $5,000 of contract service funds should be devoted each year to bringing experts in technology issues to Albany to present seminars and workshops for State agency staff.**

**When SARA develops Technology and Standards services beginning in 1988-89, the Coordinator should devote at least 10 percent time to monitoring technological developments in the area of data processing and investigating data interchange standards. In the last four years of the plan, this need will increase to approximately 40 percent time for the Coordinator to assist EDP units with**

implementing data retention guidelines and to develop tape maintenance and preservation guidelines. The estimates for resources needed by the Agency Services Bureau do not include the costs of possible capital improvements to Records Center storage vaults and environmental monitoring systems, or the costs of providing tape maintenance and preservation services.

Program Area: ARCHIVAL SERVICES

Objective 6: Develop the capacity to accession, describe, and make available for research use holdings of machine-readable records.

*Activity 17:* *Obtain data processing services and resources needed to accession machine-readable records identified as having archival value.*

Time Frame: 1988-89

Explanation: The preparation of data files for long-term storage will require access to a computer center that can provide a variety of data processing services. The computer center must have the capacity to read and process data produced in a variety of outside EDP units, to copy and reformat files, to produce printouts of selected portions of files for verification purposes, and to provide access to common statistical packages. Consultative services to assist the Archives staff with such problems as unreadable tapes, coding errors, and software dependency are also desirable.

Currently, the State Education Department's EDP Division is providing the Archives with computer access and technical support for the transfer of archival data files from two agencies to the State Archives. During FY 1988-89, the State Archives and the EDP Division will use these pilot projects to evaluate this arrangement for data processing services. These projects will also enable SARA and the EDP Division to develop more accurate estimates of the resources needed to acquire machine-readable records on a continuing basis. If the arrangement with the EDP Division is unsatisfactory, SARA will investigate alternative sources of EDP services in FY 1989-90 and select a data processing center for processing of archival data files no later than April 1, 1990.

*Activity 18:* *Develop technical specifications and procedures for the transfer of machine-readable records to the State Archives.*

Time Frame: 1988-89 to 1989-90

Explanation: During the first two years covered by this plan, the State Archives will establish standards for the transfer of archival data files from agencies to the Archives in consultation with agency records management and EDP staff. These standards will specify the storage media, densities, labeling conventions, preferred file structures, and minimum acceptable file documentation. Additional standards will be proposed later in the planning period for data created on microcomputers, in automated office systems, and for nonalphanumeric data.

*Activity 19:* *Develop processing guidelines, including appropriate levels of processing for archival data files and assign responsibility for various processing activities.*

Time Frame: 1988-89 to 1989-90

Explanation: The transfer of archival data files from agencies to the State Archives requires careful preparation of the files for long-term storage and preservation. At a minimum, a printout of selected records from each file must be compared with the documentation to ensure that codebooks and file layouts match the data provided and that coding schemes are documented adequately to allow future users to understand and interpret the data. In some cases, more elaborate file verification is needed, especially when the initial evaluation of the file uncovers problems such as missing data, coding errors, or inadequate documentation.

Verification of data and the correction of errors are labor-intensive activities. Therefore, precise guidelines regarding the appropriate level of processing for data files are needed. Such guidelines should take into account the value and uniqueness of the data and the likely costs of remedying errors in the file or documentation. Guidelines are also needed to determine which levels of staff will be assigned responsibility for verification and other accessioning procedures.

*Activity 20:* ***Accession and process data from several automated information systems each year.***

Time Frame: Ongoing

Explanation: The Preliminary Report of the Special Media Records Project and follow-up appraisal studies have identified more than 100 data files with long-term research value. In order to establish a program to preserve this data and make it available to researchers, the State Archives will accession machine-readable files from several automated information systems each year. As part of the State Archives responsibility to preserve records with enduring value, data files will be transferred to the Archives as part of the normal accessioning work flow. SARA anticipates that the volume of accessions will continue to increase as agencies develop retention and disposition schedules which include machine-readable records and as SARA staff identify files that should be transferred to the Archives for preservation. Therefore, resource requirement projections are tentative and conservative. Additional staff, supplies and services will be required if appraisal archivists identify large volumes of data files in immediate need of transfer to the Archives.

*Activity 21:* ***Evaluate descriptive standards for machine-readable records including an assessment of the applicability of the Machine-Readable Cataloging, Archives and Manuscripts Control (MARC AMC) and Machine-Readable Data File (MRDF) formats for the description of EDP records.***

Time Frame: 1988-89

Explanation: The MARC (Machine-Readable Cataloging) format is a standard format used by libraries and archives for bibliographic information. Numerous archives, including the New York State Archives, have adopted the MARC format for describing their holdings. Use of the MARC format enables repositories to enter descriptive data about their holdings into national bibliographic databases and to exchange bibliographic and other descriptive information with libraries, archives, and other repositories. Currently, seven slightly different versions of the MARC format are available for use with different types of materials, including the MARC Archives and Manuscripts Control format (AMC) and the MARC format for machine-readable data files (MRDF). Machine-readable records can be described in either or both of the formats and the archival community has not yet determined which format is preferable. During the initial year of the planning period, the staff of the State Archives will participate in an assessment of the applicability of the AMC and MRDF versions of the MARC format. This assessment will be carried out in conjunction with other data archivists through the Society of American Archivist's Task Force on Automated Records and Techniques. Following the assessment, standards and procedures will be developed for the description of EDP records in the State Archives.

*Activity 22:* ***Enter series descriptions for machine-readable records into one or more national databases as the records are processed.***

Time Frame: Ongoing

Explanation: Archival repositories are beginning to distribute and exchange information about their holdings with libraries and other repositories through national databases, such as the Research Libraries Information Network (RLIN) databases and the On-Line Computer Library Center's (OCLC) catalog. The State Archives currently is entering descriptive data about its holdings into the RLIN database. Use of such national online catalogs can increase researchers' awareness of archival holdings because researchers who might not normally contact an archives directly about holdings may discover valuable records in the course of a search at a library or another repository. This is especially true for machine-readable data files, because few potential researchers are aware that archival repositories collect and make available data in machine-readable form. Therefore, it is essential to enter descriptive information about machine-readable records in the State Archives into national databases as the records are processed.

*Activity 23:* ***Produce descriptive brochures, articles in appropriate journals, special guides, and other materials to inform potential users of the State Archives' machine-readable records holdings.***

Time Frame: Ongoing

Explanation: Special initiatives may be needed to encourage use of the State Archives machine-readable records holdings by researchers. Without publicity to the appropriate research communities, including researchers and policy analysts in State government, most potential researchers will be unaware that the State Archives acquires machine-readable records. Moreover, the types of data files acquired by the State Archives will be of interest to a wider research community than that traditionally served by the State Archives. Machine-readable records are more likely to have value for research in a wider variety of subject areas such as public administration, policy analysis, social services, criminal justice, medicine and science. Therefore, the Reference Services Unit, with assistance from the SARA External Program Office, will prepare special notices, subject brochures, and similar publicity materials to ensure that the research community is aware of the data files collected by the State Archives.

*Activity 24:* ***Meet researchers' requests as they occur.***

Time Frame: Ongoing

Explanation: Providing reference services for machine-readable records differs considerably from traditional paper files. Machine-readable data can be copied and provided to researchers at a minimal cost. Therefore, researchers can obtain data without the need to visit an archives. Often research requests are handled by providing an interested researcher with the codebook, user's guide and other descriptive materials for a file. If a data file will meet the researchers needs, arrangements are usually made to provide the researcher with a copy of the file for use at a computer center. In order to allow researchers to use the State Archives holdings, the Reference Services Unit must develop a set of access policies and procedures. Such policies should address the degree of service which the State Archives will provide, the limitations on access to data which are restricted for reasons of privacy or confidentiality, and the need, if any, to assess user's fees for copies of data files.

**Objective 7:** **Develop the capacity to store machine-readable records and perform maintenance and preservation activities for EDP records.**

*Activity 25:* ***Secure adequate storage space for one master and one back-up copy of each machine-readable data file.***

Time Frame: 1989-90

Explanation: Current estimates indicate that the State Archives will acquire approximately 180 reels of tape (master and back-up copies) over the next five years. The storage of magnetic tape requires carefully monitored and strictly controlled storage areas with as little temperature and humidity fluctuation as possible. Data archives generally maintain one master and one duplicate back-up copy as security in the event of tape deterioration. Master and back-up copies should be stored in two different locations. These special requirements should be taken into account in identifying and selecting storage areas for magnetic tapes during the next five years.

The State Archives currently is developing a long-range plan for storage of archival records in all formats. The long-range storage plan must include provisions for the storage of master and back-up copies of magnetic tapes. This plan should assess the feasibility of a shared facility with the State Records Center (See Activities 15 and 16).

### *Activity 26:* *Obtain access to data processing services necessary to perform maintenance and preservation activities.*

Time Frame: 1988-89 to 1989-90

Explanation: The long-term preservation of machine-readable records requires carefully monitoring of tape deterioration, regular maintenance routines, and occasional transfer of data to new magnetic tape. These services should be performed by a service center rather than done in-house by the State Archives staff. One criterion for selecting a data processing service center for the State Archives will be its capacity to provide tape maintenance, conservation, and copying services. The Preservation Unit will assume responsibility for monitoring these services and design a system to monitor tape maintenance activities. (See Activity 17).

### *Activity 27:* *Investigate alternative storage media.*

Time Frame: Ongoing

Explanation: Currently, magnetic tape is the standard storage medium used for the long-term retention of machine-readable data because it is less expensive and more stable than most available alternatives. However, magnetic tape is not a permanent medium and its use for storage of machine-readable records requires regular maintenance and periodic recopying to new tape. Archivists remain optimistic that new developments in storage technology may yield a more stable and cost-effective medium for the permanent retention of data in machine-readable form.

Optical disk technology may eventually provide an alternative storage medium for electronic records, although the archival properties of optical disk technology have not been fully tested. The storage of data on microfilm that can be scanned and reconverted to digital form also holds some promise. In order to take advantage of these technological developments, the Preservation Unit should monitor new developments in storage technology.

In fiscal year 1989-90, Preservation Unit staff will begin an assessment of optical disk technology for archival storage with an extensive pilot project to test the optical disk system currently under development at the National Archives in Washington, D.C. Additional resources may be requested for further development of optical disk systems at SARA, pending the results of the National Archives' pilot project.

## ARCHIVAL SERVICES: SUMMARY OF RESOURCE NEEDS

The Principal Archivist, who heads this Bureau, should spend 5 to 10 percent time coordinating and supervising the development of accessioning procedures, descriptive standards, a tape preservation and maintenance program, and reference services for EDP records. In the Collections Management Unit, the Unit Head will need 15 percent time in 1988-89 and 10 percent in 1989-90 to obtain data processing services for accessioning EDP records, establishing description guidelines and procedures, and supervising the acquisition of approximately 100 data files already identified as archival.

One additional Senior Archivist is needed in the Collections Management Unit to accession and describe machine-readable records. During FY 1988-89, SARA should fill a temporary service-funded vacancy for this activity. Beginning in FY 1989-90 this position should be transferred to permanent State funding as part of SARA's ongoing archival program. In addition, SARA will need approximately $22,000 in contract service funds each year for technical assistance, and for data processing services and supplies for accessioning machine-readable records.

The Head of the Preservation Unit should devote approximately 5 percent time in the first two years of the plan to the development of a tape preservation monitoring system and to an assessment of data processing services to ensure that services are available for tape maintenance and preservation procedures. In addition, SARA must procure adequate facilities for the storage of tapes. SARA should again request $150,000 in its 1989-90 budget for installation of vaults for storage of computer tapes and other special media records. Long-range facilities planning must include provisions for expanded storage areas for machine-readable and other special media records. Finally, the Preservation Unit should request $60,000 in funding for a project to assess the applicability of optical disk technology, beginning in FY 1989-90. This project will be carried out in conjunction with the U.S. National Archives and Records Administration which has invested over $1 million in an optical disk pilot project for archival storage. During the last three years of the plan, the Head of the Reference Services Unit should devote 5 percent time to the development of descriptive brochures and other materials which inform researchers about the availability of machine-readable records at the State Archives.

# Summary of SARA Resource Requirements

| Records Analysis and Disposition | FY 1988-89 | FY 1989-90 | FY 1990-91 | FY 1991-92 | FY 1992-93 |
|---|---|---|---|---|---|
| Personal Services (FTE) | 1.70 | 1.60 | 2.35 | 2.25 | 2.25 |
| SubTotal-Records Analysis and Disposition | $62,586 | $62,635 | $94,250 | $9[illegible],150 | $99,800 |
| **Agency Services** | | | | | |
| Personal Services (FTE) | .40 | .65 | .55 | .55 | .55 |
| Total Cost (PS) | $19,500 | $26,000 | $22,000 | $23,100 | $24,250 |
| NonPersonal Services | $5,000 | $5,800 | $5,000 | $5,000 | $5,000 |
| SubTotal-Agency Services | $24,500 | $31,800 | $27,000* | $28,100 | $29,250 |
| **Archival Services** | | | | | |
| Personal Services (FTE) | 1.35 | 1.35 | 1.05 | 1.05 | 1.05 |
| Total Cost (PS) | $40,850 | $43,420 | $34,420 | $36,220 | $38,135 |
| NonPersonal Services | $22,000 | $232,500 | $23,000 | $23,500 | $24,000 |
| Subtotal-Archival Services | $62,850 | $275,920 | $57,420 | $59,720 | $62,135 |
| **TOTAL COST** | **$149,936** | **$370,355** | **$178,670** | **$182,970** | **$191,185** |

* Does not include possible capital improvements to the State Records Center or start-up costs of a possible tape maintenance program.